Behavior Genetics
https://doi.org/10.1007/s10519-022-10127-6

ORIGINAL RESEARCH

# A General Cognitive Ability Factor for the UK Biobank

**Camille Michèle Williams**[1,5] · **Ghislaine Labouret**[1] · **Tobias Wolfram**[4] · **Hugo Peyre**[1,2,3] · **Franck Ramus**[1]

Received: 27 May 2022 / Accepted: 30 October 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

**Abstract**
UK Biobank participants do not have a high-quality measure of intelligence or polygenic scores (PGSs) of intelligence to simultaneously examine the genetic and neural underpinnings of intelligence. We created a standardized measure of general intelligence (*g* factor) relative to the UK population and estimated its quality. After running a GWAS of *g* on UK Biobank participants with a *g* factor of good quality and without neuroimaging data (N = 187,288), we derived a *g* PGS for UK Biobank participants with neuroimaging data. For individuals with at least one cognitive test, the *g* factor from eight cognitive tests (N = 501,650) explained 29% of the variance in cognitive test performance. The PGS for British individuals with neuroimaging data (N = 27,174) explained 7.6% of the variance in *g*. We provided high-quality *g* factor estimates for most UK Biobank participants and *g* factor PGSs for UK Biobank participants with neuroimaging data.

**Keywords** Intelligence · Factor analyses · GWAS · Polygenic scores · UK Biobank

## Introduction

Intelligence—our ability to learn, reason and solve problems (Arvey et al. 1994)—has been of great interest to researchers in epidemiology, neuroscience, and genetics as it predicts a wide array of educational, health, and social outcomes (e.g., Calvin et al. 2017; Deary et al. 2007; Strenze 2007). Given the numerous genetic, neural, and environmental factors that may contribute to intelligence, large-scale studies are needed to identify the respective contribution of these factors and their potential interactions on intelligence (for review see Deary et al. 2019, 2021).

The UK Biobank is an ideal database to study the causes and consequences of intelligence, with its cognitive, brain imaging, genetic, health, and environmental data on more than 500,000 British middle-aged and older adults. Yet, numerous factors make the use of cognitive tests in the UK Biobank difficult. First, not all participants completed the same number of tests and more recent tests have fewer participants (e.g., word production) because only a subset of participants returned to the test centers or participated in the online follow-up questionnaire to complete more tests when asked. Second, those who completed the same number of tests did not necessarily complete the same combination of tests. Across 501,650 participants with data on at least 1 of the 8 cognitive tests, we counted 80 different combinations of tests, with only 30,471 participants having usable data on all 8 tests. Third, the age at which a test was completed varies by test and participant, with some participants completing a test as early as 38 years old and as late as 82 years old. Fourth, some participants completed certain tests several times. Finally, some tests or similar tests with slight variations (e.g., 14 instead of 13 questions for fluid intelligence, FI) were completed at different locations: the assessment center on a touchscreen or autonomously online, with one's device.

To maximize the number of participants included in their studies on intelligence in the UK Biobank, some researchers estimated intelligence with a single test, either a verbal

Handling Editor: Chandra A. Reynolds, PhD.

✉ Camille Michèle Williams
williams.m.camille@gmail.com

1 Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Études Cognitives, École Normale Supérieure, EHESS, CNRS, PSL University, 75005 Paris, France

2 INSERM UMR 1141, Paris Diderot University, Paris, France

3 Department of Child and Adolescent Psychiatry, Robert Debré Hospital, APHP, Paris, France

4 Faculty of Sociology, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany

5 LSCP, Département d'Etudes Cognitives, École Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France

numerical reasoning (aka fluid intelligence, FI) score or a reaction time score (Davies et al. 2016; Kievit et al. 2018; Lee et al. 2018; Savage et al. 2018; Sniekers et al. 2017). Others created a general intelligence (*g*) factor from 3 to 7 cognitive variables using principal component analysis (PCA) or confirmatory factor analysis (CFA; Cox et al. 2019a, b; Hepsomali and Groeger 2021; Lyall et al. 2016; Navrady et al. 2017; de la Fuente et al. 2021). Furthermore, most studies used test scores that were neither adjusted for age nor standardized relative to a representative sample of the general population, despite the acknowledged lack of representativeness of the UK Biobank sample (Fry et al. 2017). As the UK Biobank continues to accrue data and attract new researchers, access to a standardized general factor of intelligence for most UK Biobank participants will benefit future studies that consider intelligence as a variable of interest or as a confounder.

Intelligence is moderately to highly heritable (Davies et al. 2018; Polderman et al. 2015) and individual differences in intelligence are associated with a variety of brain measures (Basten et al. 2015; Deary et al. 2010; Jung and Haier 2007). Because intelligence is highly polygenic, a person's genetic liability to being more intelligent can be quantified with a polygenic score (PGS). PGSs are calculated using a person's genotype profile and the association between each SNP and a trait quantified by genome-wide association studies (GWASs). PGS of intelligence is thought to predict 4–10.6% of the variance in intelligence (Davies et al. 2018; Hill et al. 2019; Lee et al. 2018). However, few studies examined the extent to which cerebral measures mediate the effects that intelligence PGS or the environment has on intelligence (Lett et al. 2020; Loughnan et al. 2021) because such analyses require rich datasets and GWAS results that exclude the target sample. Since the UK Biobank is consistently included in GWASs of intelligence, researchers are unable to examine the associations between intelligence PGSs and neuroanatomical, environmental, and behavioral data in the UK Biobank.

Therefore, our first aim was to create a standardized general intelligence (*g* factor) score for each UK Biobank participant with at least one cognitive test that is relative to the UK population given the participant's age, sex, and occupation. Our second aim was to create a general intelligence polygenic score (*g* PGS) for individuals with neuroimaging data in the UK Biobank, to be used by future studies aiming to link genes, brain, and intelligence.

We first standardized 8 cognitive test scores relative to the UK population and then extracted a *g* factor score with CFA from the cognitive tests for individuals that at least completed one cognitive test. We estimated the quality of the *g* factor scores of participants with missing data. We then conducted a GWAS of the *g* factor score on the UK Biobank participants with a *g* factor of good quality and without neuroimaging data (N = 187,288), and we assessed its predictive validity in the participants with neuroimaging data (N = 39,131). We assessed the external validity of our *g* factor by examining the correlation between our *g* factor and life outcomes.

This study provides cognitive measures that are partially adjusted for sampling bias in the UK Biobank and a PGS for future UK Biobank studies interested in examining the genetic associations of intelligence with neuroimaging, behavioral, and environmental measures.

## Methods

All analyses were performed in R (R Core Team 2022). Supplemental Information, Supplemental Data, and Code are anonymously available on the Open Science Framework (OSF): https://osf.io/49scv/?view_only=29e0ee6a1420461d81d234d94d549751.

The standardization of cognitive test and *g* factor scores relative to the UK population are summarized in Fig. 1.

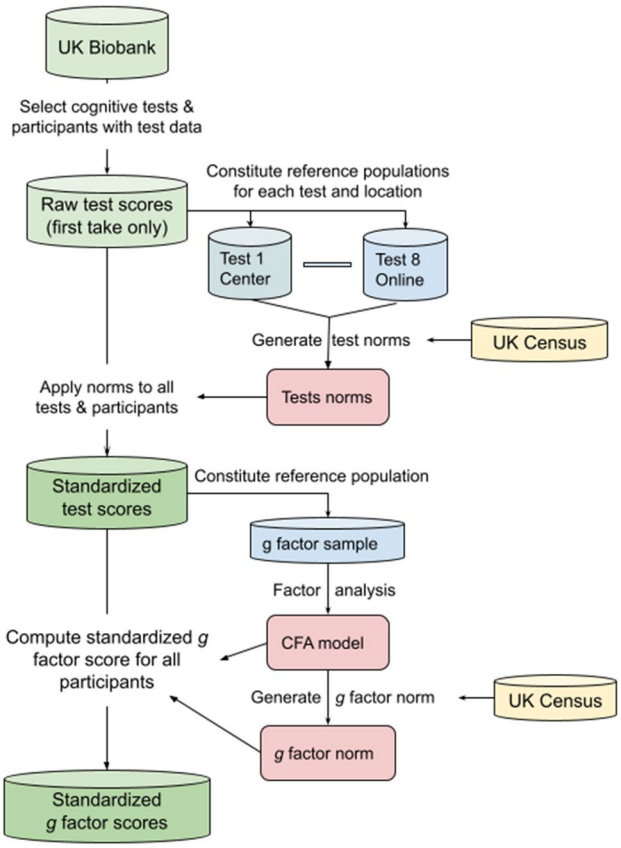


**Fig. 1** *G* factor creation pipeline. Cylinders represent datasets; inputs are UK Biobank and UK Census data (2001); blue cylinders are subsets of data (standardization samples). Boxes represent produced norms and models. Arrows represent computations (Color figure online)

## UK Biobank dataset and participants

The UK Biobank is an open-access large prospective study with phenotypic, genotypic, and neuroimaging data from more than 500,000 participants. Participants were recruited between 2006 and 2010, from the vicinity of 22 assessment centers in England, Wales, and Scotland, with an age range for inclusion of 40–69 years (Sudlow et al. 2015). Data collection continues up to date. All participants provided informed consent ("Resources tab" at https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=200). The UK Biobank received ethical approval from the Research Ethics Committee (reference 11/NW/0382) and the present study was conducted based on application 46007.

Participants performed a variety of cognitive tests, either when visiting a UK Biobank assessment center, or online during the online follow-up. Participants who did not complete any of the cognitive tests retained for this study were excluded, yielding 501,650 participants (excluded 843 participants).

The UK Biobank participants differ from the general UK population: they tend to be healthier and to have a higher socioeconomic status (e.g., more likely to own property; Fry et al. 2017; Keyes and Westreich 2019), women are overrepresented and the distribution across ages differs from the general population (ages 50–59 in 2001 overrepresented, while 30–39 are underrepresented). Therefore, we used public Census data to compensate for differences between UK Biobank participants and the UK population.

## Census data

The UK 2001 census data (Office for National Statistics 2011) was obtained from Casweb (casweb.mimas.ac.uk). We selected tables ST033 and ST034, which provide occupation categories for people currently in employment (ST033) and unemployed or economically inactive (ST034), between 16 and 74 (ST033) or 64 (ST034) years of age by sex.

We adjusted for sampling bias using the Occupational Classification in the 2001 census because occupation is correlated with intelligence (Schmidt and Hunter 2004), the data was publicly available by age, sex, and country, and job codes were similar to the UK Biobank (Standard Occupational Classification 2000—SOC2000; Office for National Statistics 2000). We did not use the 2011 census because occupation was coded using SOC2010, which differs notably from SOC2000, with no easy correspondence. We matched participants to census characteristics using their age, country, and occupation on the day the census was conducted (April 29, 2001; Supplemental Sect. 1.1). We did not use the level of qualification because it was not publicly available by age, country, and sex in the 2001 or 2011 UK Census. We examined the distribution of standardized and age-adjusted test scores with *cNorm* (Sect. 2.3.4) at the center (Fig. S1) and online (Fig. S2) across job codes.

## Cognitive tests

### Test selection

UK Biobank participants could complete several cognitive tests every time they visited the UK Biobank assessment centers (category 100026) and during the online follow-up (category 116). We used eight cognitive tests to create the *g* factor (bolded tests in Table 1). Some participants completed some tests several times. We only considered the first occurrence of each test to best reflect the stable part of general intelligence, before aging and cognitive decline.

### Obtaining raw scores

To obtain a raw score for each test, we had to select between variables when several measures were provided for a test and/or transform these measures. We excluded participants with abnormal results (e.g., too many errors in the Symbol digit substitution test, indicating non-compliance with the test instructions) or who did not finish the test. Retained measures, transformations, and exclusion criteria are described in Table 2.

### Standardization of test scores

Standardization served two purposes: (1) to adjust for age effects (since intellectual performance varies with age), and (2) to provide a test score relative to the UK population (Fig. 1). We created a common norming model for males and females. We simultaneously performed two adjustments:

(1) An age adjustment by using the semiparametric continuous norming method was proposed by Lenhard et al. (2016). With this method, raw scores are modeled as a function of both standard scores and an explanatory variable, age when taking the test in this case.
(2) A socio-demographic adjustment: by using standardization samples and computing weights to apply to participants, to compensate for the socio-demographic differences between the UK Biobank population and the complete UK population.

To do so, we first created standardization samples for each test and location (online/center), with about 32,000 to 497,000 participants. Details regarding the standardization sample creation and the number of participants in each sample are in Supplemental Sect. 1.2.

We then used cell weighting to adjust measures from the standardization samples to reflect the UK population

**Table 1** UK Biobank cognitive tests considered for this study

| Test (included tests in bold) | UKB links (C: center, O: online) | Description | Number of participants | Included or excluded |
|---|---|---|---|---|
| **FI: fluid intelligence** | C<br>O 118 | Under a time limit of 2 min, answer a set of 13 (center) or 14 (online) numerical and verbal reasoning questions | C = 205,333<br>O = 123,613 | Included |
| **MAT: matrix pattern completion** | C 501 | Select the element that best completes matrix pattern blocks. 15 Puzzles | C = 33,657 | Included |
| **TWR: tower rearranging** | C 503 | Looking at an illustration of three pegs (towers), on which three differently-colored hoops have been placed, find how many moves it would take to rearrange the hoops into another specific position. 18 Puzzles | C = 33,381 | Included |
| **MEMN: numeric memory** | C 100029<br>O 120 | Memorize 2 digits displayed on the screen. After they disappear for 3 s, enter them. Every time a sequence is correctly remembered, the next sequence is made one digit longer, up to a maximum of 12 digits | C = 82,865<br>O = 111,062 | Included |
| **MEMS: pairs matching** | C 100030<br>O 117 | Memorize the position of matching pairs of cards. Once the cards are turned face down, find as many pairs as possible in the fewest tries. Up to 3 rounds, with an increasing number of pairs (3, 6, 8) | C = 498,730<br>O = 118,528 | Included |
| *MEMW: paired associate learning* | C 506 | Memorize 12 pairs of words shown for 30 s in total. After an interval (different test), see the first word of 10 of these pairs and select the matching second word from 4 alternatives | C = 34,045 | Excluded: ceiling effect |
| *MEMP: prospective memory* | C 100031 | Early in the test session, the participant is shown "At the end of the games we will show you four colored shapes and ask you to touch the Blue Square. However, to test your memory, we want you to touch the Orange Circle instead" | C = 211,952 | Excluded: only 1 question |
| *MEML: lights pattern memory* | C 100028 | See pictures of houses which have some windows lit. After a 10-s delay, indicate which windows were lit | C = 3714 (pilot only) | Excluded: too few participants |
| **RT: reaction time** | C 100032 | Watch two cards on the screen. If they are the same, press a button-box as quickly as possible | C = 496,829 | Included |
| **SDS: symbol digit substitution** | C 502<br>O 122 | Identify the digits attached to each symbol in a grid, by using another grid linking symbols to digits as a key | C = 33,679<br>O = 118,466 | Included |

**Table 1** (continued)

| Test (included tests in bold) | UKB links (C: center, O: online) | Description | Number of participants | Included or excluded |
|---|---|---|---|---|
| **TMT: trail making** | C 505<br>O 121 | Click sequentially on a set of digits in circles scattered around the screen (numeric path), then on a set of digits/letters (alphanumeric path) | C = 34,045<br>O = 104,028 | Included |
| *VOC: picture vocabulary* | C 504 | Indicate which of 4 images is most closely related to a displayed word. Difficulty varies according to the correctness of the previous answers | C = 33,606 | Excluded: view-only field. Data is currently not available in May 2022 |
| *WRD: word production* | C 100077 | State as many words beginning with the letter 'S' as possible within 1 min | C = 3744 (pilot only) | Excluded: too few participants |

The numbers of participants are taken from UK Biobank's showcase, across all instances, and include uncompleted tests

**Table 2** Raw scores transformations for the included cognitive tests

| Test | Measures used | Raw score computation |
|---|---|---|
| FI: fluid intelligence | Number of correct answers [0–14]<br>Fields 20016 (center) and 20191 (online) | Measure unchanged |
| MAT: matrix pattern completion | Number of correct answers [0–15]<br>Field 6373 (center only) | Measure unchanged |
| TWR: tower rearranging | Number of correct answers [0–18]<br>Field 21004 (center only) | Measure unchanged |
| MEMN: numeric memory | Maximum number of digits remembered correctly [0–12]<br>Fields 4282 (center) and 20240 (online) | Measure unchanged |
| MEMS: pairs matching | Numbers of correct and incorrect matches in each round. The test has up to 3 rounds, with increasing difficulty (more pairs to remember). Access to a round is subject to a high score in the previous round<br>Fields 10136/398 (center) and 20131 (online)<br>Fields 10137/399 (center) and 20132 (online) | Score computed as follows:<br>– Each correct pair earns 2 points in rounds 1 and 2, 1 point in round 3<br>– Each incorrect pair loses 1 point<br>– Within each round, negative scores are brought back to zero |
| SDS: symbol digit substitution | Number of correct matches and number of attempts<br>Fields 23324 (center) and 20159 (online) Fields 23323 (center) and 20195 (online) | Score = the number of correct matches<br>Exclusion criteria: scores with more than 35 attempts and less than 65% correct matches (participants likely did not follow test instructions, for example by repeatedly entering the same digit). These outlier limits were computed as ± 3SD from the mean |
| RT: reaction time | Mean time to correctly identify matches<br>Field 20023 (center only) | Score = − log(mean time to correctly identify matches)<br>Higher scores represent better performance<br>No exclusion criteria because response times out of the 50 ms to 2000 ms range were already excluded |
| TMT: trail making | Duration to complete alphanumeric path trail<br>Fields 6350 (center) and 20157 (online) | Score = − log(duration to complete alphanumeric path trail)<br>Higher scores represent better performance<br>This measure had higher correlations with other tests than a measure based on a combination of the numeric and alphanumeric trails |

characteristics: For each standardization sample, we computed the proportion of participants for each possible combination of country, sex, age range at census, occupation status, and occupation SOC group. We compared the proportion of UK Biobank participants in each cell to the 2001 census and created weights for each cell by dividing the census proportion by the UK Biobank proportion. See Supplemental Sect. 1.3 for details and an example.

We used the cNorm package (Lenhard et al. 2018) to compute norming models on the standardization samples with census weights, using the semiparametric continuous norming method. We modeled raw scores as a function of standard scores (percentiles) and age at test completion. Age at test completion is provided in field 21003 for tests taken at the assessment center and in fields 20134 to 20138 for tests taken online. This age differs from the age used to compute census weighting factors, which is the participant's age on the day of the 2001 census.

We applied the norming models to the whole dataset and obtained standardized test scores for all participants on the tests they took (Fig. 1).

## *G* factor

We created a *g* factor score for all participants who completed at least one of the eight cognitive tests using CFA. The *g* factor was standardized relative to the UK population. We also evaluated the impact of missing test scores on the quality of the *g* factor.

### CFA parameters

We performed a CFA with one-factor loading on the eight cognitive tests. We estimated the CFA model with the *lavaan* R package (Rosseel 2012). We used the full information maximum likelihood (FIML) estimator to make use of all data points even for cases with missing values, estimated the mean structure, and set the variance of the latent variable to 1 to estimate each observed variable loading. Model fit was assessed using commonly used model fit indices: the Tucker–Lewis index (TLI), the comparative fit index (CFI), standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA). Good fit was established with a CFI and TLI > 0.95, a RMSEA < 0.06 and a SRMR < 0.08 (Hu and Bentler 1999). See Supplemental Sect. 1.4 for a discussion on the choice of factor analysis.

### G factor score standardization

We created a standardization sample with 496,990 participants who had data for the census variables: sex, age on census day, and occupation on census day (countries were merged, see Supplement Sect. 1.3), to compute census weighting factors for these participants (Fig. 1).

We computed factor scores using the regression estimation method, which maximizes validity (DiStefano et al. 2009). We then computed these scores' weighted mean and standard deviation, using the census weighting factors. We subtracted the weighted sample mean from the raw factor scores and divided the result by the weighted sample standard deviation, to obtain factor scores with a general population mean of 0 and a standard deviation of 1.

### Evaluation of the *g* factor score quality in the presence of missing data

We examined how well a *g* factor score computed using a subset of tests (called partial factor score) correlates with the factor score that would have been obtained from the full set of eight tests (called full factor score) by looking at the correlation between the full and the partial factor scores for each the 80 subsets of tests present in the data, in the 30,471 participants who completed all the tests.

## Analyses

In the following analyses, we included participants whose combination of cognitive tests allowed for a correlation with the complete *g* factor of 0.70 or higher (N = 261,701). This threshold was chosen to maximize the robustness of the factor as well as the number of participants for which we would generate a *g* factor.

### Testing validity: correlations with alternative estimates of *g* and life outcomes

To examine the external validity of our *g* measure by examining correlations in complete cases between our *g* factor and FI, life and health outcomes expected to correlate with intelligence [e.g., educational attainment (EA), income, deprivation indices, etc., life and health outcomes described in Supplemental Data S8]. In brief, we selected well-being, household income before taxes, highest qualification as well as the Townsend deprivation score—a deprivation score of an individual's postal code from the census data—and the index of multiple deprivations that regroups several deprivation indices which vary by country. The latter include subindices such as health, income, education, employment, and housing. The index of multiple deprivations and its subindices come from a UK government qualitative study of deprived areas in British local councils and are calculated separately for England, Wales, and Scotland. Common multiple deprivation scores across countries were combined into a single variable for the correlation matrix (Supplemental

Data S8). We adjusted each measure for sex and age at which the measure was reported.

### Genetic analyses

We conducted genetic analyses to create PGSs for individuals with neuroimaging data for future UK Biobank studies. A detailed overview of the genetic analyses in the main text and the supplements are available in Supplemental Sect. 1.6 In the supplements, we compared our GWASs results to those from previous GWAS studies on intelligence, cognitive performance, or EA with public data to ensure that the lead SNPs impacted genes previously associated with intelligence, cognitive performance, or EA. We additionally report genomic loci from the full GWAS that were not previously associated with cognition using functional mapping and annotation of GWASs (FUMA; http://fuma.ctglab.nl/; Watanabe et al. 2017; Supplemental Sect. 1.6.4).

**Main analyses** In brief, we conducted a *g* factor GWAS on 187,288 individuals without neuroimaging data or twins/siblings with neuroimaging data (No Neuroimaging GWAS). We removed participants with neuroimaging data and their siblings to maintain the independence of predictions and prevent overfitting. We controlled for sex, center, chip, birth year, and the first 40 principal components (PCs) of the genotyped data. All GWASs were conducted on 5,319,661 variants with a linear mixed model-based association analysis using a sparse genetic relationship matrix to control for relatedness (fastGWA; Jiang et al. 2019) from the genome-wide complex trait analysis (GCTA) package (Yang et al. 2016).

Using sBayesR (Lloyd-Jones et al. 2019), we created PGSs from the summary statistics of the No Neuroimaging GWAS for individuals with either neuroimaging data or siblings with neuroimaging data to assess the predictive power of genetic variance from the No Neuroimaging GWAS on the *g* factor (for details see Supplemental Sect. 1.6.5.1). After excluding individuals from non-British ancestry, first or second-degree cousins, and parent–offsprings, we adjusted the *g* factor PGS for sex, birth year, and the first 40 PCs and then examined its association with *g*.

**Additional analyses** We conducted additional analyses to answer the following questions.

(1) ***Does the g factor PGS explain more variance in g than the Fluid Intelligence (FI) PGS?*** Using the GWAS and PGS procedures described above, we ran a GWAS of FI on 180,722 individuals without neuroimaging data or twins/siblings with neuroimaging data (FI GWAS) and calculated FI PGS for individuals with neuroimaging data with the GWAS and PGS parameters described above. After excluding individuals from non-British ancestry, first or second-degree cousins, and parent–offsprings and adjusting the FI PGS for sex, birth year, and the first 40 PCs, we examined the percentage of variance explained in *g* and FI, separately.

(2) ***Once you control for between-family factors, what proportion of variance in g does the g PGS predict?*** To quantify to what extent the polygenic signal captures genetic effects that pass through the environment (indirect genetic effects or genetic nurture; Howe et al. 2021), we conducted family fixed-effects analyses. To do so, we first ran a *g* factor GWAS on the sample from the No Neuroimaging GWAS without siblings (no family, No Neuroimaging GWAS). We used the summary statistics from this GWAS to create PGSs for individuals with siblings. After excluding individuals from non-British ancestry, first or second-degree cousins, we adjusted the PGS for sex, birth year, and the first 40 PCs. Finally, we ran the family fixed effects model with and without including sibship as a random effect and reported the change in explained variance of the PGS on the *g* factor when adjusting or not for sibling pairs (for details see Supplemental Sect. 1.6.5.2).

(3) ***Do our g factor and FI measures have similar genetic correlations and heritabilities as the Cognitive Performance and Educational Attainment (EA) measures from*** Lee et al. (2018)***?*** We examined whether the genetic influences underlying our *g* factor and FI measure were similar to the genetic influences underlying the EA (i.e., years of education, or achieved educational level) and cognitive performance (i.e., measured as FI in the UK Biobank and with a *g* factor in the COGENT and CHARGE consortiums) reported by Lee et al. (2018). We compared our findings to the Lee et al. (2018) results because this is the largest genetic study of cognition to date and results are publicly available. To do so, we calculated the genetic correlations between our *g*-factor GWAS summary statistics, our FI GWAS summary statistics, and the publicly available EA and *g* factor summary statistics. We additionally calculated the heritability estimates of each summary statistics file to examine whether our *g* factor was more heritable than our FI measure, previous *g* factor measures, and EA. These analyses were conducted with linkage disequilibrium score regression using the ldsc function from the GSEM package (Grotzinger et al. 2019).

(4) ***If we include participants with a poor g factor measure (r < 0.7), are the genetic correlation and heritabilities similar?*** Using the GWAS described above, we conducted a GWAS on 307,009 individuals with a *g* factor measure and without neuroimaging data or twins/siblings with neuroimaging data (low quality No Neuroimaging *g* factor GWAS). If the genetic correla-

tion and heritabilities were similar or better for this low *g* quality GWAS compared to the No Neuroimaging GWAS, we should be able to use this larger sample size to obtain more lead SNPs and better polygenic predictions.

#### Testing validity: correlations with alternative estimates of *g*, life outcomes, neuroimaging, and genetic measures

We conducted additional correlational analyses on a subset of participants that had neuroimaging data and PGSs. We examined correlations on complete cases between our *g* factor and alternative measures of intelligence for the UK Biobank (FI alone and *g* factors with a subset of the tests we used), life and health outcomes expected to correlate with intelligence (e.g., EA, income, deprivation indices, etc., life and health outcomes in Supplemental Data S8 on OSF), total brain volume (TBV; Williams et al. 2021), and the FI and *g* factor PGS in individuals with neuroimaging data.

We compared our *g* factor score to alternate measures of intelligence by transforming the cognitive variables and extracting the *g* factor as done by Cox et al. (2019a, b) and de Nooij et al. (2020), which used a different combination of cognitive tests and factoring methods with a similar sample to the one we used to calculate PGSs. The authors used cognitive tests completed at the center during the neuroimaging visit (Instance 2) and included tests that were not initially available at the first center visit. Cox et al. (2019a, b) created a latent factor using CFA from the MAT, the SDS, the FI, and the TMTB cognitive tests, and de Nooij et al. (2020) extracted the first PC of the numerical memory, the FI, the SDS, the TMTB, the MAT, and the TWR cognitive tests.

Since our *g* factor is adjusted for age, we created alternative *g* factors that are adjusted for age in the CFA or after extracting the first PC. We additionally similarly controlled for sex to examine whether differences between our *g* factor and the *g* alternatives could be explained by sex differences.

## Results

### Cognitive tests

We compared the distribution of the standard test scores before and after adjusting for the difference between the UK Biobank population and the general UK population (Fig. 2). After adjusting for age and census weights, the distribution of the scores shifted to the right, indicating a relatively higher score in the UK Biobank relative to the UK norm (Supplemental Sect. 2.1.1). In some cases, the distribution was not normal because of its categorical nature (e.g., MEMN) or because of threshold effects (e.g., MEMS). Correlations between standard scores ranged from 0.07 (MEMN and RT) to 0.50 (TMT and SDS). TMT and FI had the highest correlation coefficients with other cognitive tests (Fig. S3; Supplemental Sect. 2.1.2).

### CFA results

#### Model fit

The CFA model fit on 501,650 participants (30,307 with complete data; N Women = 272,955; N Men = 228,695) was good (CFI = 0.955, TLI = 0.938, RMSEA = 0.024, SRMR = 0.028). The *g* factor accounted for 29% of the variance across cognitive tests and the loadings ranged from 0.77 (TMT) to 0.277 (RT) (Fig. 3; Supplemental Sect. 2.1.3). Sex differences in cognitive and *g* factor scores are available in Supplemental Sect. 2.1.4.

#### Distribution of the *g* factor before and after census correction

We compared the distribution of the *g* factor scores before (M = −0.004, SD = 0.993) and after (M = 0.086, SD = 1.001) adjusting for the difference between the UK Biobank population and the general UK population (d = 0.09). After adjusting, the factor score distribution shifted to the right, indicating a relatively higher score in the UK Biobank relative to the UK norm (Fig. 4; *g* distribution by job category in Supplemental Sect. 2.1.5).

#### G factor quality

The quality of the *g* factor and the number of individuals for each possible combination of completed tests are available in Supplemental Sect. 2.1.6. For example, if we select participants with any of the first 73 cognitive test combinations observed in the UK Biobank, the worst factor scores of these participants would have a correlation of 0.70 with the ideal, 8-tests factor scores, and the number of available participants will be over 261,701.

### Analysis of individuals with a *g* factor quality over 0.70

#### Correlations: alternative *g* factors and life outcomes

We examined the external validity of our *g* measure by examining the association in complete cases between our *g* factor measure, FI (which is often used as a proxy of *g* in most studies), and life outcomes collected at the first center visit or online to estimate correlation coefficients on a larger number of participants (Fig. 5).

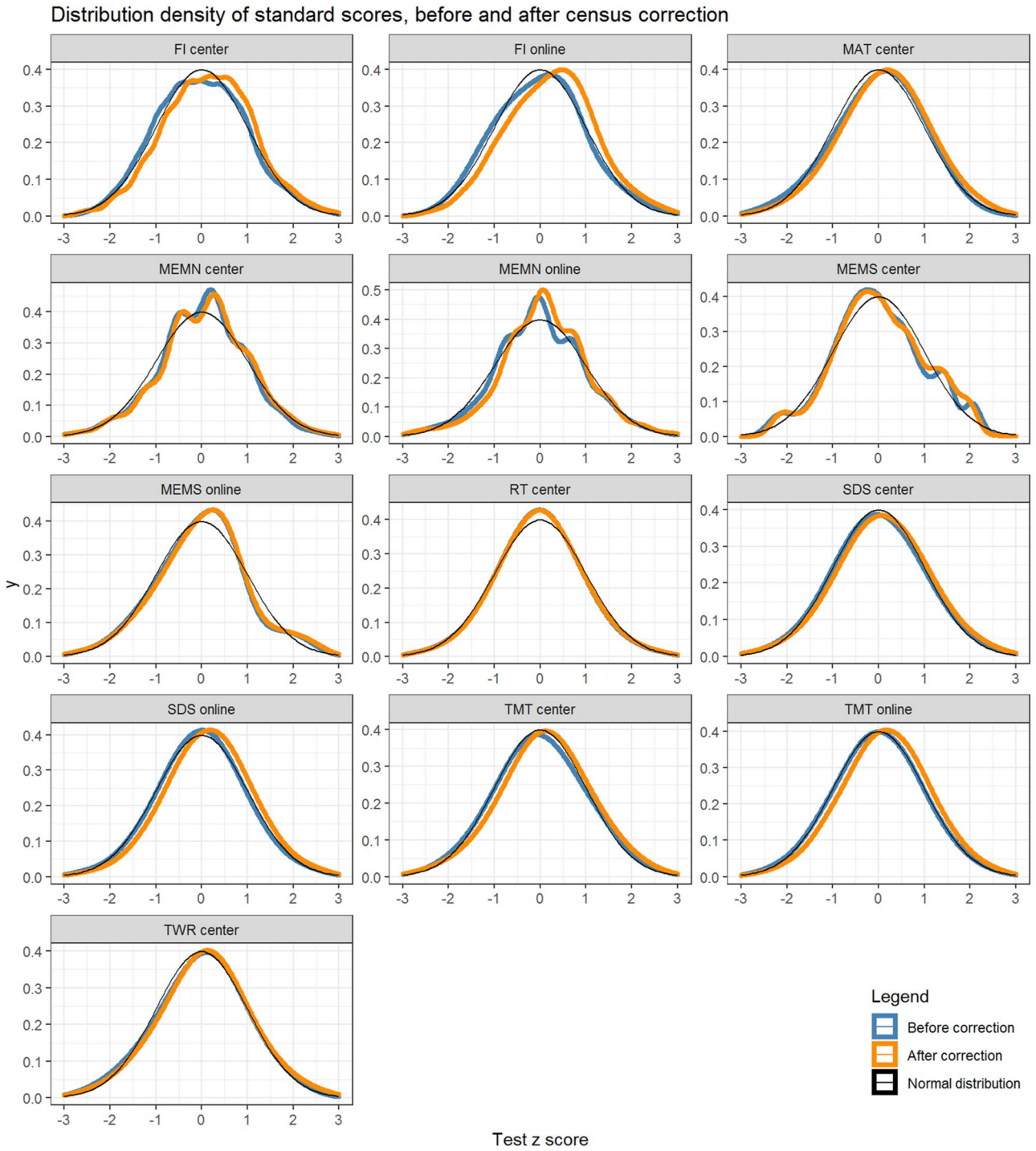


**Fig. 2** The distribution of cognitive tests before (blue) and after (orange) correcting for sociodemographic differences between the UK Biobank population and the UK population (2001 census). *FI* fluid intelligence, *RT* reaction time, *MAT* matrix pattern completion, *TWR* tower rearranging, *MEMN* numeric memory, *MEMS* pair matching, *SDS* symbol digit substitution, *TMT* trail making (Color figure online)

## Genetic analyses

**Main results** We identified 150 approximately lead SNPs attaining genome-wide significance in the full sample GWAS (h2 = 0.201, SE = 0.008) and 100 in the No Neuroimaging Sample GWAS (h2 = 0.197, SE = 0.008). There were 127 genomic risk loci associated with the *g* factor in the full GWAS and 84 genomic risk loci with the *g* factor in

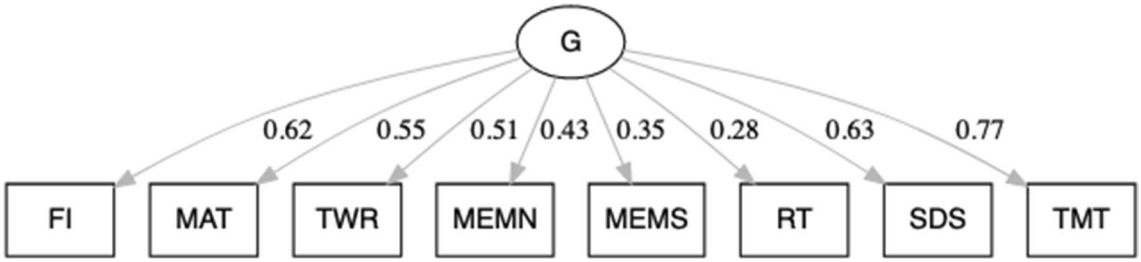


**Fig. 3** Confirmatory factor analysis of UK Biobank cognitive tests. Analyses were conducted with full information maximum likelihood with the lavaan package (Rosseel 2012). Explained variance 29%. *FI* fluid intelligence, *MAT* matrix pattern completion, *TWR* tower rearranging, *MEMN* numeric memory, *MEMS* pair matching, *SDS* symbol digit substitution, *TMT* trail making. Loadings from completely standardized solutions (i.e., standardized observed and latent variables)

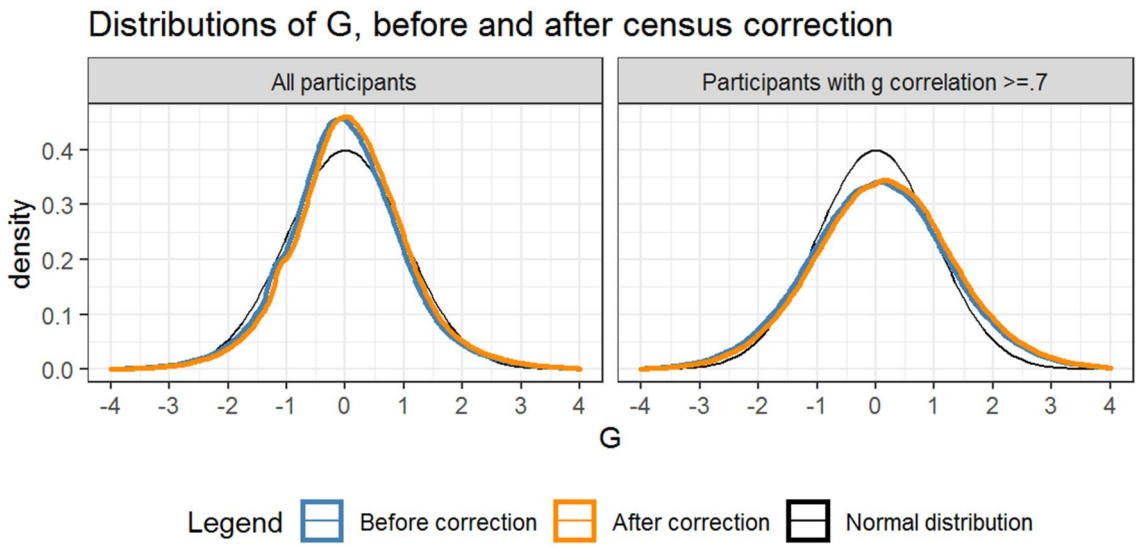


**Fig. 4** The distribution of the *g* factor scores before and after census correction for all participants (left) and a subset of participants (right). The subset of participants had a *g* factor score from a combination of subtests that allowed for a minimum correlation of 0.70 between the partial *g* factor score and the full *g* factor score

the No Neuroimaging GWAS ($P < 5 \times 10^{-8}$; Supplemental Data S1–S2 on OSF). See Fig. S5 for the Manhattan and QQ-plots. Sixty-three lead SNPs from the full GWAS were previously associated with genes known to impact intelligence, cognitive performance, or EA (Supplemental Sects. 2.2.3. and 2.2.4). A highly similar polygenic signal was captured by the *g*-factor from the No Neuroimaging GWAS, the full GWAS, and the publicly available GWASs of intelligence (rg = 0.90–1; Fig. S10; Lee et al. 2018; Savage et al. 2018).

From the No Neuroimaging GWAS, we created *g* factor PGS for 38,866 individuals and a FI PGS for 38,642 individuals who either had neuroimaging data or siblings with neuroimaging data and *g* factor quality greater than $r \geq 0.70$. Of the 39,866 individuals with neuroimaging data, 23,689 had a *g* factor quality of r = 1 (Table S9).

After excluding individuals from non-British ancestry and first or second-degree cousins and parent–offspring, we adjusted the *g* and FI values for sex, year of birth, and the first 40 genetic PCs. The *g* factor PGS created from the No Neuroimaging GWAS explained 7.6% of the variance in the *g* factor of 26,082 individuals with neuroimaging data and 1092 twins or siblings without neuroimaging data (N = 27,174). The FI created from the No Neuroimaging GWAS explained 6.6% of the variance in FI of individuals with either neuroimaging data or siblings with neuroimaging data (N = 26,360; Supplemental Sect. 2.2.5).

## Additional analyses

(1) ***Does the g factor PGS explain more variance in g than the FI PGS?*** The *g* factor PGS created from the No Neuroimaging GWAS explained 1.6% more vari-

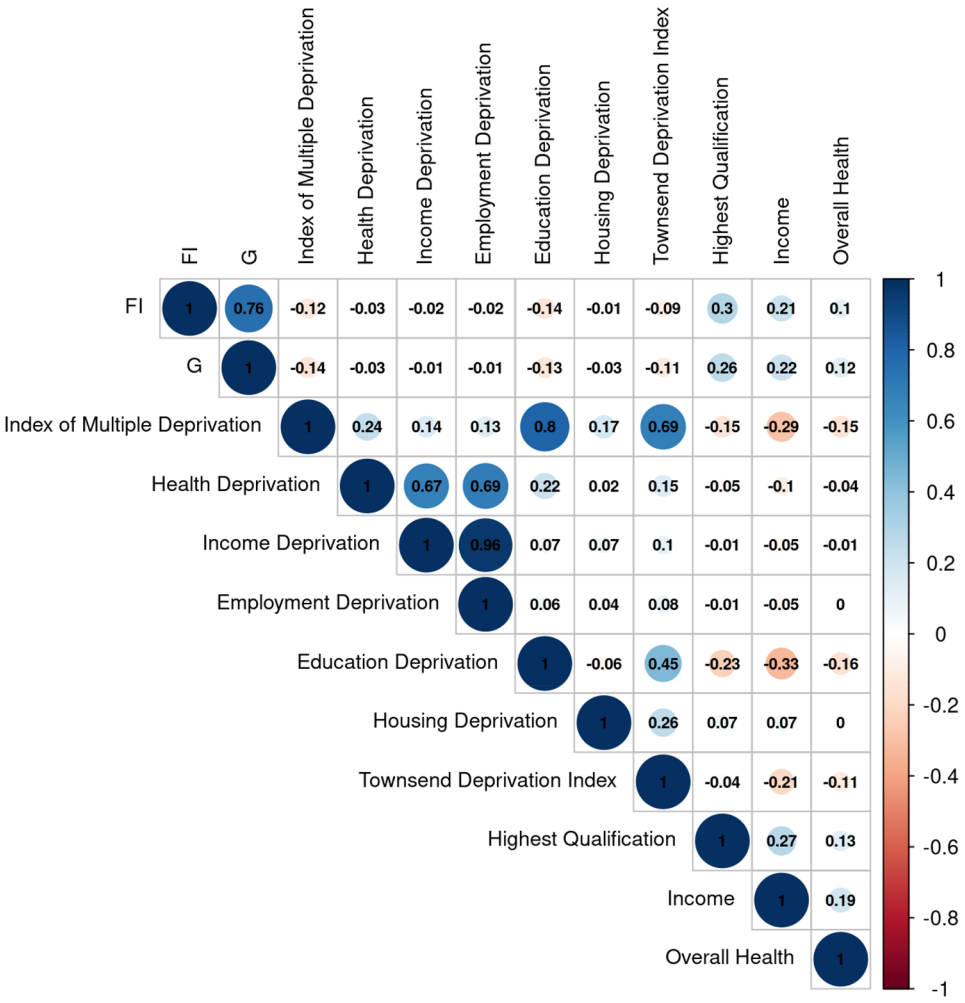


**Fig. 5** Correlation between age and sex-adjusted *g* factor Scores and health and life outcomes. Pearson correlation coefficients were estimated on 181,327 individuals without missing data. All measures are adjusted for age at which the measure was taken and sex. G corresponds to the *g* factor of individuals with a combination of cognitive tests that allowed for a correlation of 0.70 or higher between their actual *g* factor and what their *g* factor would have been if they had completed all tests. *FI* fluid intelligence

ance in *g* than the FI PGS created from the No Neuroimaging GWAS, which explained 6.0% of the variance in the *g* factor (N = 26,360). The *g* factor PGS created from the No Neuroimaging GWAS explained 5.8% of the variance in FI (N = 26,360).

(2) ***Once you control for between-family factors, what proportion of variance in g does the g PGS predict?*** After excluding individuals from non-British ancestry and first or second-degree cousins and correcting for sex, year of birth, and the first 40 genetic PCs, the family fixed-effect analysis on the *g* factor PGS from the No Family GWAS summary statistics showed a decrease in explained variance 7.5% from to 4.2% and a reduction of 10% in the effect size (N = 14,601). See Supplemental Sect. 2.2.4 for additional PGS analyses.

(3) ***Do our g factor and FI measures have similar genetic correlations and heritabilities as the g factor and Educational Attainment (EA) measures from*** Lee et al. (2018)***?*** The SNP heritability of the present *g* factor was similar to the heritability of Lee's *g* factor (h2 = 0.199, SE = 0.008) and greater than Lee's EA (h2 = 0.151, SE = 0.004; Tables S10, S11). Lee's *g* factor was highly correlated to our various estimates of the *g* factor 0.92–0.93, suggesting that they capture common genetic influences (Fig. S10).

(4) ***Do participants with a low-quality g factor estimate (r < 0.7) impact the results?*** We found that the heritability of the Low-Quality No Neuroimaging GWAS (h2 = 0.127, SE = 0.005) was much lower than the heritability of the No Neuroimaging GWAS (h2 = 0.197, SE = 0.008; Table S10) and that the genetic correlation between the Low-Quality No Neuroimaging GWAS and the No Neuroimaging GWAS was of 0.98 (Fig. S10), suggesting that although both GWASs are measuring overlapping genetic effects, the Low-Quality No Neuroimaging GWAS has more measurement error. The genetic correlations of the Low-Quality No Neuroimaging GWAS were lower with Lee's *g* factor (0.87 vs. 0.93) and EA (0.5 vs. 0.55), suggesting that doubling

the sample size by including *g* factor estimates of lower quality is counterproductive for the GWAS.

### Correlations: alternative *g* factors, life outcomes, neuroimaging, and genetic measures

We conducted separate correlation analyses on our *g* factor with the alternative *g* factors, neuroimaging, and PGSs because these data were only available for participants who visited the center for the neuroimaging visit.

The alternative *g* factors were created from cognitive test scores from their neuroimaging visit and cognitive tests that were only available at the second visit. The correlation between our study's *g* factor (adjusted for age, not sex) and the cognitive tests ranged from 0.25 (RT) to 0.84 (TMT), whereas the age-adjusted *g* factors from previous studies ranged from about 0.16 (RT) to 0.71 (MAT; Fig. S13).

Correlation coefficients between our study's *g* factor and alternative *g* factors were high (0.85 to 0.89; Fig. S13). The alternative *g* factor measures were more correlated to the highest qualification achieved (r = 0.26–0.29 vs. 0.24) and income (r = 0.18–0.19 vs. 0.16) than the present *g* factor because the cognitive tests included in the alternative *g* factors were the cognitive tests with the highest correlations with highest qualification and income (Fig. S15).

The *g* factor from the present study and the *g* factors calculated as done by previous studies were highly correlated after adjusting for the age of test completion and sex. The *g* factors positively correlated with the PGS, highest qualifications, income before tax, overall health, and TBV, and negatively with the deprivation indices (Fig. 6). We also looked at the correlation coefficients of the *g* factors with well-being and additional deprivation indices, which reflect the degree of housing, employment, education, etc. deprivation in an area. However, these correlations were small and therefore reported in Fig. S16.

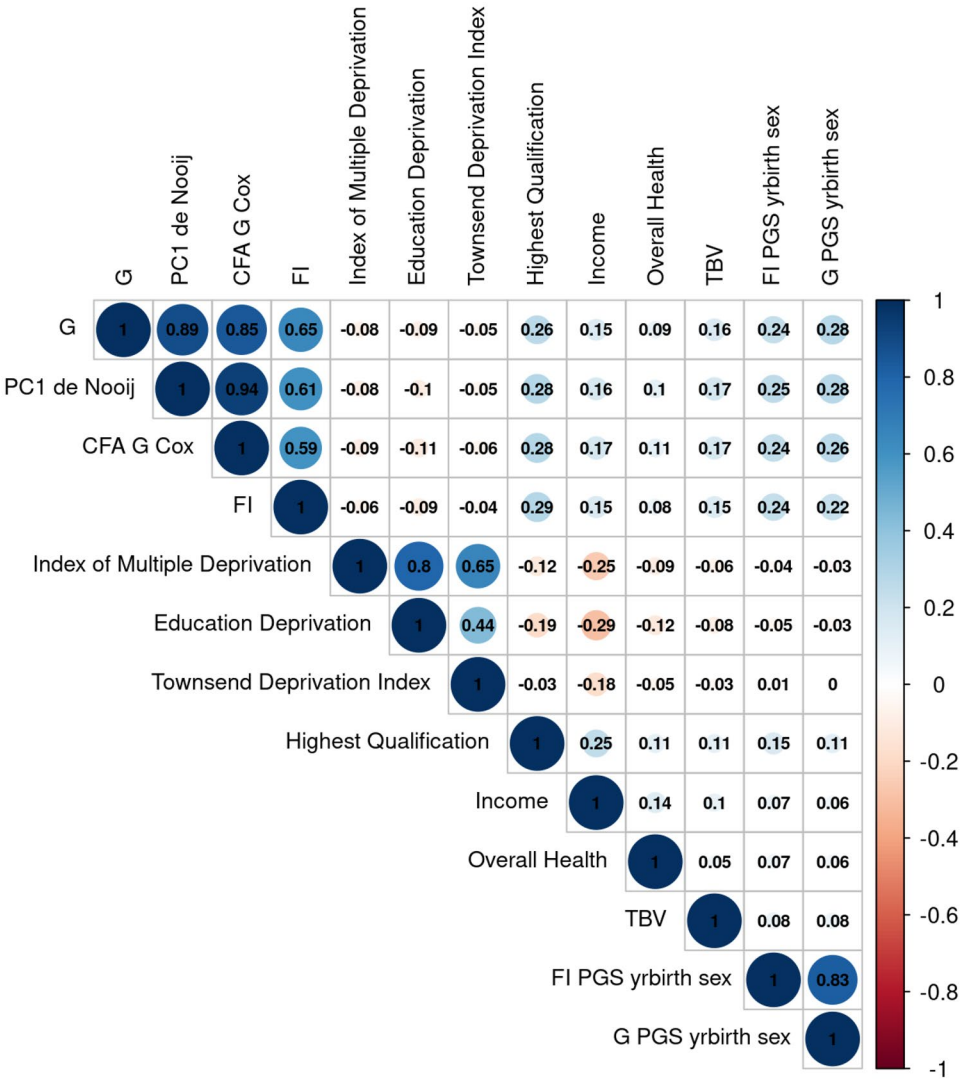


**Fig. 6** Correlation between age and sex-adjusted G factor scores, selected life and health outcomes, total brain volume (TBV), and sex and year of birth adjusted polygenic scores (PGSs). Pearson correlation coefficients were estimated on 13,085 British individuals without missing data without first- or second-degree cousins and parent–offspring. We included all life and health variables with an r > or = 0.1 with *g* except for the Townsend deprivation index. *CFA* confirmatory factor analysis, *PC1* 1st principal component, *G* general factor for intelligence. G corresponds to the *g* factor of individuals with a combination of cognitive tests that allowed for a correlation of 0.70 or higher between their actual *g* factor and what their *g* factor would have been if they had completed all tests. *FI* fluid intelligence, *Income* income before tax. PGSs were adjusted for sex and birth year (year birth)

## Discussion

We aimed to create an age-standardized *g* factor measure that is relative to the UK population and provide a PGS for UK Biobank participants with neuroimaging data. Unlike previous studies on the *g* factor in the UK Biobank, we partially adjust for sampling bias on the *g* factor in the UK Biobank, we provide an estimation of the *g* factor's quality for each participant with missing data, and we perform a *g* factor GWAS excluding participants with neuroimaging data to allow future studies on the genetic, environmental, and neural correlates of intelligence in the UK Biobank. Our *g* factor was highly correlated with alternative *g* factor measures of intelligence in the UK Biobank and their correlations with life and health outcomes were similar. The *g* factor PGS of 26,082 UK Biobank individuals with neuroimaging data and 1092 siblings without neuroimaging data explained 7.6% of the variance in the intelligence score.

The *g* factor from the present study and alternative *g* factor measures were similarly correlated to life outcomes, such as household income before tax or highest qualification (i.e., level of education) when adjusting for age and sex. The slightly higher correlations between the alternative *g* factor measures and the highest qualification achieved compared to our *g* factor could be explained by the tests used to create the alternative *g* factors, which correlated the most with the highest qualification. Therefore, some outcomes, such as the highest qualification, may not be as highly correlated to general intelligence as previously thought but may be more correlated with specific cognitive tests.

We report a negative correlation between our *g* factor and the Townsend index, an SES measure that reflects a person's material deprivation based on unemployment rates, non-car ownership, non-homeownership, and household overcrowding in their postal code. This negative correlation coincides with the negative correlation between neighborhood deprivation and EA (Garner and Raudenbush 1991), as well as previous findings that childhood IQ remains stable across old age (Deary et al. 2000), predicts later SES outcomes (Deary et al. 2005). We additionally found a negative correlation between our *g* factor and the index of multiple deprivation of the area where a person resides, which was largely explained by the negative correlation between the *g* factor and the education deprivation index of the area where a person resides. The education deprivation index was measured by a score reflecting child and adolescent school performance (e.g., English, math, and science exams Stage 3 exams) and adult skills (e.g., the proportion of adults with no or low qualifications) in a given geographical area.

The *g* PGS created from the No Neuroimaging GWAS summary statistics explained 7.7% of the variance in the *g* factor of individuals with neuroimaging data and siblings with neuroimaging data. The largest GWAS study of cognition to date (Lee et al. 2018) similarly found that their Cognitive Performance PGS explained 7% of the variance in Cognitive Performance for individuals in the Wisconsin Longitudinal Study, a study that used cognitive tests with similar properties to their discovery GWAS. The *g*-factor in our study captured the same polygenic signal as the *g*-factor of Lee et al. (2018; rg = 0.92). However, they found that the Cognitive Performance PGS from the summary statistics of their multi-trait analysis GWAS (MTAG) of CP yielded more significant SNPs and thus, explained 9.7% of the variance in CP for individuals in the Wisconsin Longitudinal Study. Although the MTAG summary statistics may explain more variance in *g*, these summary statistics cannot be applied to the UK Biobank individuals with neuroimaging data because they were included in the discovery GWAS.

The FI PGS explained 1.6% less variance in *g* than the PGS of the *g* factor, suggesting that the *g* PGS is a better genetic predictor of *g*. The genetic correlations between FI and the *g* factor were high, suggesting that FI and *g* have very similar genetic influences. The *g* factor and FI PGSs also similarly correlated with TBV, the most correlated brain measure to intelligence (Deary et al. 2021). Taken together, our findings suggest that a common genetic component to FI and the *g* factor may explain the genetic association between intelligence and TBV. Although the FI PGS may be sufficient when investigating global brain size associations with intelligence, the *g* factor PGS still explains a larger range of variance in general intelligence than FI and its use should be favored when controlling for the genetic components of intelligence.

The present study has several limitations. First, the selection of cognitive tests currently available in the UK Biobank severely underrepresents verbal ability. Only two of the eight cognitive tests included in this study were verbal: the FI test (verbal-numerical reasoning) and the numeric memory test (digit span). The UK Biobank is currently in the process of adding verbal tests to the cognitive assessment (e.g., picture vocabulary and word production), which would justify the calculation of a new *g* factor once they are completed by a sufficient number of participants.

The underrepresentation of verbal skills may partly explain why women had a slightly lower *g* factor than men (d = −0.13). Another non-exclusive explanation may be that male and female UK populations are unequally sampled in the UK Biobank, with women representing 54.4% of the entire sample. Thus, women with lower general intelligence may have been oversampled compared to men with lower intelligence. Finally, participants were born between 1934 and 1971 (median 1950), a period when women in the UK may have had inferior educational opportunities, preventing

them from reaching their intellectual potential. The UK Biobank may thus not be suitable to reliably estimate sex differences in cognitive abilities and other phenotypes associated with cognitive ability.

Second, although the loadings of the present study correspond to those previously reported by UK Biobank studies (Cox et al. 2019a, b; de la Fuente et al. 2021), the highest loadings on the *g* factor in the UK Biobank differ from those reported across studies using psychometric tests to measure intelligence. Specifically, the Raven's progressive matrices test is expected to have the highest loading on *g* across psychometric tests around 0.7 (Gignac 2015; Gignac and Watkins 2013). And yet, we and other UK Biobank studies report the highest loading for trail-making and FI (Cox et al. 2019a, b; de la Fuente et al. 2021) and a loading of around 0.5 for the matrices. One study examined the concurrent validity of each UK Biobank cognitive test by reporting the correlation between each UK Biobank cognitive test and one to several well-validated standard cognitive tests of the same cognitive domain (reference tests). The authors found that the UK Biobank TMT B strongly correlated at 0.66 with the reference TMT B test and that the UK Biobank matrices correlated at 0.57 with the reference matrices test. Overall, they concluded that the UK Biobank tests load strongly on general cognitive ability. They additionally measured test–retest reliability after a mean of 28 days and report that the test–retest reliability was greater than 0.5 but that the mean performance on some tests, such as FI increased at Time 2. Considering that we took the first instance of test completion, our *g* factor should not be prone to repeat testing effects (Fawns-Ritchie and Deary 2020).

We were limited when correcting for the socio-demographic imbalance in the UK Biobank. Due to occupation coding constraints, we had to use the 2001 census data, instead of the 2011 census data, and we only determined the occupation on census day of 71.5% of participants. Moreover, we were limited by the number of variables on which we adjusted the UK Biobank sample as we did not have access to the qualification data by age, sex, and country. Although further variables should be adjusted to provide a *g* factor that is perfectly relative to the UK population, we nonetheless provide cognitive test scores and a *g* factor measure that are age-standardized and, to some extent, adjusted for sampling bias in the UK Biobank.

Finally, the *g* factor score was calculated from different subsets of tests and although we took the first instance of a test, some tests were taken at different ages. Therefore, although we attempted to provide a *g* factor measure that most resembles pre-aging adult intelligence, some test scores may already be influenced by cognitive decline. However, one study examining test–retest reliability over a period of 4 years, which may reflect cognitive decline, reported that most UK Biobank cognitive tests show reasonable stability except for the visual memory task (pairs-matching; Lyall et al. 2016) and another study found declines in cognitive abilities before 65 years of age were small (Cornelis et al. 2019). Therefore, considering that the majority of participants were 64 years or younger and that the median age was 60, our *g* factor measure likely reflects pre-cognitive-decline intelligence scores.

The present study provides cognitive test scores and a *g* factor score for UK Biobank participants that are adjusted for age and partially adjusted for sampling bias as well as a *g* PGS for UK Biobank individuals with neuroimaging data that explained 7.6% of the variance in the *g* factor. The behavioral and genetic scores from this study will enable the simultaneous investigation of the associations between the brain, genes, and intelligence, which are currently rare in the present literature (Deary et al. 2021). Taken together, the present study offers robust measures of intelligence that will foster homogeneity in intelligence research within the UK Biobank and provides summary statistics and PGSs for future studies interested in examining the genetic associations of intelligence with neuroimaging, behavioral, and environmental measures.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10519-022-10127-6.

**Funding** This work received support under the Program "Investissements d'Avenir" launched by the French Government and implemented by l'Agence Nationale de la recherche (ANR) with the Reference ANR-17-EURE-0017 and ANR-10-IDEX-0001-02 PSL. This research has been conducted using the UK Biobank Resource. Declarations of interest: none.

**Data availability** This research has been conducted using data from UK Biobank, a major biomedical database (http://www.ukbiobank.ac.uk/). Restrictions apply to the availability of these data, which were used under license for this study: application 46007. Supplements and code available on OSF: https://osf.io/49scv/?view_only=29e0ee6a1420461d81d234d94d549751.

**Code availability** This research has been conducted using data from UK Biobank, a major biomedical database (http://www.ukbiobank.ac.uk/). Restrictions apply to the availability of these data, which were used under license for this study: application 46007. Supplements and code available on OSF: https://osf.io/49scv/?view_only=29e0ee6a1420461d81d234d94d549751.

## Declarations

**Conflict of interest** Camille Michèle Williams, Ghislaine Labouret, Tobias Wolfram, Hugo Peyre and Franck Ramus states that there is no conflict of interest.

**Ethical approval** The UK Biobank received ethical approval from the Research Ethics Committee (reference 11/NW/0382) and the present study was conducted based on application 46 007.

**Informed consent** All participants provided informed consent ("Resources tab" at https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=200).

## References


Arvey RD, Bouchard TJ, Carroll JB, Cattell RB, Cohen DB, Dawis RV, Willerman L (1994) Mainstream science on intelligence. Wall Str J 13(1):18–25

Basten U, Hilger K, Fiebach CJ (2015) Where smart brains are different: a quantitative meta-analysis of functional and structural brain imaging studies on intelligence. Intelligence 51:10–27. https://doi.org/10.1016/j.intell.2015.04.009

Calvin CM, Batty GD, Der G, Brett CE, Taylor A, Pattie A, Čukić I, Deary IJ (2017) Childhood intelligence in relation to major causes of death in 68 year follow-up: Prospective population study. BMJ. https://doi.org/10.1136/bmj.j2708

Cornelis MC, Wang Y, Holland T, Agarwal P, Weintraub S, Morris MC (2019) Age and cognitive decline in the UK Biobank. PLoS ONE 14(3):e0213948. https://doi.org/10.1371/journal.pone.0213948

Cox SR, Ritchie SJ, Fawns-Ritchie C, Tucker-Drob EM, Deary IJ (2019a) Structural brain imaging correlates of general intelligence in UK Biobank. Intelligence. https://doi.org/10.1016/j.intell.2019.101376

Cox SR, Ritchie SJ, Fawns-Ritchie C, Tucker-Drob EM, Deary IJ (2019b) Brain imaging correlates of general intelligence in UK Biobank. BioRxiv. https://doi.org/10.1101/599472

Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaars SP, Harris SE, Ritchie SJ, Luciano M, Fawns-Ritchie C, Lyall D, Cullen B, Cox SR, Hayward C, Porteous DJ, Evans J, McIntosh AM, Gallacher J, Craddock N, Pell JP et al (2016) Genome-wide association study of cognitive functions and educational attainment in UK Biobank (*N*=112,151). Mol Psychiatry 21(6):758–767. https://doi.org/10.1038/mp.2016.45

Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, Hagenaars SP, Ritchie SJ, Marioni RE, Fawns-Ritchie C, Liewald DCM, Okely JA, Ahola-Olli AV, Barnes CLK, Bertram L, Bis JC, Burdick KE, Christoforou A, DeRosse P et al (2018) Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun 9(1):2098. https://doi.org/10.1038/s41467-018-04362-x

Deary IJ, Whalley LJ, Lemmon H, Crawford JR, Starr JM (2000) The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish mental survey. Intelligence 28(1):49–55. https://doi.org/10.1016/S0160-2896(99)00031-8

Deary IJ, Taylor MD, Hart CL, Wilson V, Smith GD, Blane D, Starr JM (2005) Intergenerational social mobility and mid-life status attainment: influences of childhood intelligence, childhood social factors, and education. Intelligence 33(5):455–472. https://doi.org/10.1016/j.intell.2005.06.003

Deary IJ, Strand S, Smith P, Fernandes C (2007) Intelligence and educational achievement. Intelligence 35(1):13–21. https://doi.org/10.1016/j.intell.2006.02.001

Deary IJ, Penke L, Johnson W (2010) The neuroscience of human intelligence differences. Nat Rev Neurosci 11(3):201–211. https://doi.org/10.1038/nrn2793

Deary IJ, Harris SE, Hill WD (2019) What genome-wide association studies reveal about the association between intelligence and physical health, illness, and mortality. Curr Opin Psychol 27:6–12. https://doi.org/10.1016/j.copsyc.2018.07.005

Deary IJ, Cox SR, Hill WD (2021) Genetic variation, brain, and intelligence differences. Mol Psychiatry. https://doi.org/10.1038/s41380-021-01027-y

de la Fuente J, Davies G, Grotzinger AD, Tucker-Drob EM, Deary IJ (2021) A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. Nat Hum Behav 5(1):49–58. https://doi.org/10.1038/s41562-020-00936-2

de Nooij L, Harris MA, Adams MJ, Clarke T-K, Shen X, Cox SR, McIntosh AM, Whalley HC (2020) Cognitive functioning and lifetime major depressive disorder in UK Biobank. Eur Psychiatry. https://doi.org/10.1192/j.eurpsy.2020.24

DiStefano C, Zhu M, Mindrila D (2009) Understanding and using factor scores: considerations for the applied researcher. Pract Assess Res Eval 14(1):20

Fawns-Ritchie C, Deary IJ (2020) Reliability and validity of the UK Biobank cognitive tests. PLoS ONE 15(4):e0231627. https://doi.org/10.1371/journal.pone.0231627

Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R, Allen NE (2017) Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. Am J Epidemiol 186(9):1026–1034. https://doi.org/10.1093/aje/kwx246

Garner CL, Raudenbush SW (1991) Neighborhood effects on educational attainment: a multilevel analysis. Sociol Educ 64(4):251–262. https://doi.org/10.2307/2112706

Gignac GE (2015) Raven's is not a pure measure of general intelligence: implications for g factor theory and the brief measurement of g. Intelligence 52:71–79. https://doi.org/10.1016/j.intell.2015.07.006

Gignac GE, Watkins MW (2013) Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. Multivar Behav Res 48(5):639–662. https://doi.org/10.1080/00273171.2013.804398

Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, Ip HF, Marioni RE, McIntosh AM, Deary IJ, Koellinger PD, Harden KP, Nivard MG, Tucker-Drob EM (2019) Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav 3(5):513–525. https://doi.org/10.1038/s41562-019-0566-x

Hepsomali P, Groeger JA (2021) Diet and general cognitive ability in the UK Biobank dataset. Sci Rep 11(1):11786. https://doi.org/10.1038/s41598-021-91259-3

Hill WD, Marioni RE, Maghzian O, Ritchie SJ, Hagenaars SP, McIntosh AM, Gale CR, Davies G, Deary IJ (2019) A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Mol Psychiatry 24(2):169–181. https://doi.org/10.1038/s41380-017-0001-5

Howe LJ, Nivard MG, Morris TT, Hansen AF, Rasheed H, Cho Y, Chittoor G, Lind PA, Palviainen T, van der Zee MD, Cheesman R, Mangino M, Wang Y, Li S, Klaric L, Ratliff SM, Bielak LF, Nygaard M, Reynolds CA et al (2021) Within-sibship GWAS improve estimates of direct genetic effects, p. 2021.03.05.433935. https://doi.org/10.1101/2021.03.05.433935

Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 6(1):1–55. https://doi.org/10.1080/10705519909540118

Jiang L, Zheng Z, Qi T, Kemper KE, Wray NR, Visscher PM, Yang J (2019) A resource-efficient tool for mixed model association analysis of large-scale data. Nat Genet 51(12):1749–1755. https://doi.org/10.1038/s41588-019-0530-8

Jung RE, Haier RJ (2007) The Pareto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. Behav Brain Sci 30(2):135–154; discussion 154–187. https://doi.org/10.1017/S0140525X07001185

Keyes KM, Westreich D (2019) UK Biobank, big data, and the consequences of non-representativeness. Lancet 393(10178):1297. https://doi.org/10.1016/S0140-6736(18)33067-8

Kievit RA, Fuhrmann D, Borgeest GS, Simpson-Kent IL, Henson RNA (2018) The neural determinants of age-related changes in fluid intelligence: a pre-registered, longitudinal analysis in UK Biobank. Wellcome Open Res. https://doi.org/10.12688/wellcomeopenres.14241.2

Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M, Nguyen-Viet TA, Bowers P, Sidorenko J, Linnér RK, Fontana MA, Kundu T, Lee C, Li H, Li R, Royer R, Timshel PN, Walters RK, Willoughby EA et al (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat Genet 50(8):1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Lenhard A, Lenhard W, Suggate S, Segerer R (2016) A continuous solution to the norming problem. Assessment 25(1):112–125. https://doi.org/10.1177/1073191116656437

Lenhard A, Lenhard W, Gary S (2018) CNORM—generating continuous test norms. https://doi.org/10.13140/RG.2.2.25821.26082

Lett TA, Vogel BO, Ripke S, Wackerhagen C, Erk S, Awasthi S, Trubetskoy V, Brandl EJ, Mohnke S, Veer IM, Nöthen MM, Rietschel M, Degenhardt F, Romanczuk-Seiferth N, Witt SH, Banaschewski T, Bokde ALW, Büchel C, Quinlan EB et al (2020) Cortical surfaces mediate the relationship between polygenic scores for intelligence and general intelligence. Cereb Cortex 30(4):2708–2719. https://doi.org/10.1093/cercor/bhz270

Lloyd-Jones LR, Zeng J, Sidorenko J, Yengo L, Moser G, Kemper KE, Wang H, Zheng Z, Magi R, Esko T, Metspalu A, Wray NR, Goddard ME, Yang J, Visscher PM (2019) Improved polygenic prediction by Bayesian multiple regression on summary statistics. Nat Commun 10(1):5086. https://doi.org/10.1038/s41467-019-12653-0

Loughnan RJ, Palmer CE, Thompson WK, Dale AM, Jernigan TL, Fan CC (2021) Gene-experience correlation during cognitive development: evidence from the Adolescent Brain Cognitive Development (ABCD) StudySM. bioRxiv. https://doi.org/10.1101/637512

Lyall DM, Cullen B, Allerhand M, Smith DJ, Mackay D, Evans J, Anderson J, Fawns-Ritchie C, McIntosh AM, Deary IJ, Pell JP (2016) Cognitive test scores in UK Biobank: data reduction in 480,416 participants and longitudinal stability in 20,346 participants. PLoS ONE 11(4):e0154222. https://doi.org/10.1371/journal.pone.0154222

Navrady LB, Ritchie SJ, Chan SWY, Kerr DM, Adams MJ, Hawkins EH, Porteous D, Deary IJ, Gale CR, Batty GD, McIntosh AM (2017) Intelligence and neuroticism in relation to depression and psychological distress: evidence from two large population cohorts. Eur Psychiatry J Assoc Eur Psychiatr 43:58–65. https://doi.org/10.1016/j.eurpsy.2016.12.012

Polderman TJC, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, Posthuma D (2015) Meta-analysis of the heritability of human traits based on fifty years of twin studies. Nat Genet 47(7):702–709. https://doi.org/10.1038/ng.3285

R Core Team (2022) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rosseel Y (2012) lavaan: an R package for structural equation modeling. J Stat Softw 48(2):1–36

Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, Nagel M, Awasthi S, Barr PB, Coleman JRI, Grasby KL, Hammerschlag AR, Kaminski JA, Karlsson R, Krapohl E, Lam M, Nygaard M, Reynolds CA, Trampush JW et al (2018) Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet 50(7):912–919. https://doi.org/10.1038/s41588-018-0152-6

Schmidt FL, Hunter J (2004) General mental ability in the world of work: occupational attainment and job performance. J Personal Soc Psychol 86(1):162–173. https://doi.org/10.1037/0022-3514.86.1.162

Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E, Taskesen E, Hammerschlag AR, Okbay A, Zabaneh D, Amin N, Breen G, Cesarini D, Chabris CF, Iacono WG, Ikram MA, Johannesson M, Koellinger P, Lee JJ et al (2017) Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat Genet 49(7):1107–1112. https://doi.org/10.1038/ng.3869

Strenze T (2007) Intelligence and socioeconomic success: a meta-analytic review of longitudinal research. Intelligence 35(5):401–426. https://doi.org/10.1016/j.intell.2006.09.004

Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R (2015) UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 12(3):e1001779. https://doi.org/10.1371/journal.pmed.1001779

Watanabe K, Taskesen E, Van Bochoven A, Posthuma D (2017) Functional mapping and annotation of genetic associations with FUMA. Nat Commun 8(1):1–11

Williams CM, Peyre H, Toro R, Ramus F (2021) Neuroanatomical norms in the UK Biobank: the impact of allometric scaling, sex, and age. Hum Brain Mapp 42(14):4623–4642. https://doi.org/10.1002/hbm.25572

Yang J, Lee SH, Wray NR, Goddard ME, Visscher PM (2016) GCTA-GREML accounts for linkage disequilibrium when estimating genetic variance from genome-wide SNPs. Proc Natl Acad Sci USA 113(32):E4579–E4580. https://doi.org/10.1073/pnas.1602743113

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.